Bayesian response adaptive randomization

Shengping Yang PhD, Gilbert Berdine MD

I am planning a randomized trial to compare five different treatment strategies for breast cancer. To generate sufficient evidence for identifying the most effective treatment, an adequate number of patients will need to be enrolled in each arm. To improve efficiency and potentially reduce costs, I am considering using a design that incorporates Bayesian response adaptive randomization. Is this a viable and appropriate option for such a multi-arm trial?

1. Introduction

Randomization has long been the cornerstone of clinical trial design, serving to minimize bias and ensure comparability between treatment groups. The classic randomized controlled trial (RCT), typically with fixed allocation ratios (e.g., 1:1 or 1:2), is widely regarded as the gold standard for evaluating therapeutic efficacy.[1] However, this design has important limitations, particularly when multiple treatment arms are involved, that a large number of participants is often required, and the design does not allow for adaptation based on accumulating evidence during the course of the trial.

Over the past two decades, the field of adaptive trial design has evolved in response to growing demands for more flexible, efficient, and ethically sensitive approaches. Adaptive designs allow for pre-specified modifications to key trial parameters, such as sample size, allocation ratio, or number of treatment arms, based on interim analyses, without compromising the scientific validity or integrity of the study.[2]

One of the most ethically appealing and statistically innovative approaches is response adaptive randomization, in which the probability of assigning a patient to a particular treatment arm changes over time in response to observed outcomes. While frequentist methods can also support adaptive randomization, Bayesian methods are particularly well-suited due to their capacity to incorporate prior knowledge and continuously update treatment effect estimates as new data accrue.

Bayesian Response Adaptive Randomization (Bayesian RAR) thus represents a fusion of Bayesian statistical modeling and patient-centered ethics. It aims to allocate more patients to better-performing treatments as evidence emerges, while preserving statistical rigor. This approach has gained particular traction in complex and rapidly evolving settings, such as precision medicine, where flexibility, efficiency, and real-time decision-making are essential.

2. Bayesian RAR

Bayesian RAR uses Bayes' theorem to update the probability distribution of treatment effects as new data accumulates. The resulting posterior probabilities of success for each treatment arm are then calculated and used to guide the randomization scheme – assigning patients with higher probability to treatments that demonstrate a greater likelihood of success based on trial outcomes (e.g., treatment response).[3]

2.1 Outcome Assessment

2.1.1 Outcome type

Clinical trials often involve a variety of outcome types, including continuous, binary, count-based, and time-to-event (survival) outcomes, as well as outcomes measured at different time points, i.e., short-term versus long-term.

While Bayesian RAR can, in theory, accommodate all these outcome types, trials with short-term or early outcomes, for example, tumor response at week 3, are particularly well suited to this approach. These outcomes become available soon after treatment and provide timely data for updating posterior probabilities

Corresponding author: **Shengping Yang**
Contact Information: Shengping.Yang@pbrc.edu
DOI: 10.12746/swjm.v13i56.1531

and adjusting randomization probabilities (or for early stopping decisions; see Section 2.2 for details). Access to such early indicators of efficacy enables the adaptive design to function effectively within the trial's timeline.

In contrast, longer-term outcomes, such as progression-free survival (PFS) or overall survival (OS), are less practical for guiding adaptive allocation due to their delayed availability. Depending solely on these outcomes can slow or even halt adaptation, thus compromising the primary advantage of RAR. To mitigate this limitation, many trials employ intermediate or surrogate endpoints that are more immediately observable and reasonably predictive of long-term success. For instance, tumor response measured by response evaluation criteria in solid tumors (RECIST) is frequently used as a surrogate for overall survival in oncology. While not a perfect substitute, it strikes a practical balance between biological relevance and feasibility for adaptive decision-making.

2.1.2 Multiple outcomes

Some trials collect more than one clinically relevant outcome, for example, tumor response and toxicity, or response and progression. In such cases, Bayesian models can simultaneously incorporate multiple endpoints, enabling a more nuanced evaluation of treatment performance, as well as decision on whether to stop an arm during an interim analysis.

2.2 Key Components of Bayesian RAR

Bayesian RAR is grounded in the core principles of Bayesian inference. Its implementation requires several interconnected components that work together to guide real-time decisions about treatment allocation and trial conduct.

2.2.1 Prior distribution

The prior distribution reflects the state of belief about treatment effects before any trial data are observed. This can be:

- Non-informative (or weakly informative), where little to no prior knowledge is assumed, allowing the data to dominate the inference.
- Informative, incorporating data from previous trials, meta-analyses, or expert clinical judgment. The choice of prior, which is often conjugate for easy computation, has implications for how quickly the model adapts and how early conclusions can be drawn. In early-phase exploratory trials, informative priors may be useful to maximize learning, while confirmatory trials often prefer conservative priors to avoid bias.

2.2.2 Likelihood function

The likelihood function quantifies the probability of observing the data, given assumed treatment effects. It forms the bridge between observed outcomes and the underlying model. The likelihood varies depending on the type of outcome (binary, continuous, count, survival) and the statistical model used.

2.2.3 Posterior distribution

Using Bayes' theorem, the posterior distribution combines the prior and the likelihood to yield an updated probability distribution over treatment effects, conditioned on the observed data. This distribution provides the foundation for all adaptive decisions, including,

- Treatment allocation: Randomization probabilities are derived based on the posterior probability that each treatment (often compared with a control arm) is the most effective.
- Stopping decisions (details in section 2.2.5): A treatment arm may be dropped if its posterior probability of being optimal falls below a predefined threshold, or if its probability of unacceptable toxicity exceeds a specified limit.
- Trial termination (details in section 2.2.6): The entire trial may be stopped early if the posterior probability that one treatment is superior reaches a predefined threshold (e.g., >99%), indicating that further randomization is unlikely to change the conclusion.

The calculations of posterior distribution, as well as the probability of each treatment being the most effective treatment, compared to the control arm, is mathematically involved and thus are not the focus of this article. The goal of these decision making is to conserve

resources and accelerate access to effective treatment while maintaining statistical rigor.

2.2.4 Randomization function

The randomization function translates posterior probabilities into treatment assignment probabilities to strike a balance between exploration (learning about all treatments) and exploitation (favoring the most effective treatment).

When multiple outcomes (e.g., efficacy and safety) are present, a composite utility function can be used to derive a net benefit score. Posterior probabilities are then computed based on this utility, guiding treatment allocation accordingly. As the trial progresses and more outcomes are observed, these probabilities are updated dynamically, and randomization probabilities adapt in real time.

2.2.5 Stopping rules

Bayesian RAR incorporates stopping rules to drop treatment arms for efficacy, futility, or safety, based on accumulating evidence:[4]

- Futility: An arm may be dropped if its probability of being the best is very low (e.g., <5%). For example, if Arm C has only a 2% posterior probability of being the most effective treatment after 100 enrolled patients, then arm C may be discontinued.
- Safety: If an arm is associated with unacceptable toxicity, for example, the posterior probability of unacceptable toxicity exceeds a high threshold (e.g., >95%), then it may be stopped regardless of efficacy.

These rules enhance the ethical and scientific rigor of the trial by ensuring that adaptation is not only guided by positive responses but also by early signals of failure or harm. Therefore, early stopping can enhance efficiency and reduce unnecessary patient exposure while preserving scientific validity.

2.2.6 Early trial termination

A Bayesian RAR trial may be terminated early when sufficient evidence accumulates to support a clear conclusion about treatment superiority or equivalence. This can happen when:

- Overwhelming efficacy

  If a treatment achieves a high predefined posterior probability of being the best treatment after enrolling, for example, 150 patients, which is well before the planned maximum sample size, then the trial may be stopped, and results reported early.
- All other arms are dropped due to futility or safety concerns.
- A predefined maximum level of certainty has been reached regarding treatment comparisons.

2.3 Interim vs continuous posterior probability calculation

Bayesian updating process is typically implemented either at predefined interim analyses or continuously as outcome data become available. In trials with scheduled interim analyses, after an initial run-in period, the posterior distribution of treatment effects is estimated for each arm. From these distributions, the posterior probability that each treatment is the most effective is calculated. These same probabilities are then used to guide the randomization of new participants – assigning them with higher probability to treatments with greater chances of success – until the next interim analysis is performed.

Alternatively, in a fully adaptive or continuous updating approach, posterior probabilities are calculated for each new participant using the most current outcome data from all previously enrolled patients. This allows the randomization scheme to respond more rapidly to emerging evidence, potentially improving both trial efficiency and participant outcomes. However, this approach requires efficient and timely logistical execution, including real-time data collection, rapid outcome assessment, and immediate updating of randomization algorithms.

2.4 An illustrative example

Consider a five-arm Bayesian RAR trial evaluating four experimental neoadjuvant therapies (arms A–D) against a standard control (arm E) in patients with

triple-negative breast cancer. The primary endpoint is tumor reduction at week 6.

The trial begins with equal randomization (20% per arm). After enrolling the first 100 patients, a Bayesian interim analysis estimates the effect size of each treatment relative to the control and calculates the posterior probability that each arm is the most effective, for example: A (22%), B (62%), C (15%), D (1%). Based on these probabilities, randomization weights are updated to favor more promising treatments: more patients are assigned to Arm B, while Arm D is dropped for futility (pre-defined as posterior probability $\leq$1%). Other criteria for dropping an arm, such as safety concerns (e.g., excessive toxicity), may also apply. The control arm is maintained with the same allocation probability as the most effective experimental arm, using normalized probabilities.

This adaptive strategy is repeated at the second and third interim analyses. At each stage, posterior probabilities are re-estimated to guide allocation and determine whether any arm should be dropped or declared superior. At the end of the study, the most effective experimental arm will be formally compared with the control to assess whether it provides a statistically significant improvement in tumor reduction.

This Bayesian RAR design enables dynamic learning and ethical trial conduct by minimizing patient exposure to ineffective treatments and increasing assignment to more promising options. Critically, it also allows for early trial termination if there is strong evidence of either efficacy (i.e., one treatment clearly superior) or futility (i.e., no meaningful treatment effect). By incorporating adaptive decision-making, this design can substantially reduce total sample size, shorten trial duration, and lower costs, while maintaining statistical validity. The ability to update treatment allocation in real time based on emerging data improves trial efficiency and supports more patient-centered, precision-driven clinical research.

3. Advantages of Bayesian RAR

Compared to fixed allocation ratio randomization, Bayesian RAR offers several advantages:

- Ethical appeal: Patients are more likely to receive better-performing treatments as the trial progresses, addressing ethical concerns about exposing participants to suboptimal therapies.[5]
- Statistical efficiency: Bayesian RAR can lead to smaller expected sample sizes, particularly when treatment effects are large or differences become apparent early.[6]
- Flexibility: Bayesian methods naturally handle uncertainty, missing data, and complex endpoint structures, including time-to-event outcomes and hierarchical models.
- Incorporation of prior knowledge: Prior data or expert opinion can be formally integrated, improving trial efficiency and reducing redundancy.
- Decision-theoretic extensions: Bayesian RAR can be embedded within broader decision-making frameworks, such as adaptive stopping rules or seamless phase II/III designs.[7]

4. Challenges and Limitations

Despite its appeal, Bayesian RAR comes with practical and conceptual hurdles:

- Operational complexity:

  One of the biggest challenges of Bayesian RAR is that its implementation requires real-time data capture, rapid statistical analysis, and robust infrastructure to ensure correct allocation in a timely manner.

- Limited number of treatment arms:

  When a trial includes only a small number of arms, particularly in the case of two-arm trials, Bayesian RAR may offer limited advantages.[8] In such settings, a fixed allocation ratio (e.g., 1:1) is often more statistically efficient, especially when outcome variances are similar across arms. The strengths of RAR, such as preferential allocation to superior treatments and early discontinuation of underperforming arms, are more impactful in multi-arm trials.

- Regulatory acceptance:

  While Bayesian designs are gaining traction, regulators may scrutinize the choice of priors, simulation-based operating characteristics, and adaptive rules more rigorously than for classical trials.

- Delayed outcomes:

  RAR depends on timely outcome assessments to adjust probability for treatment allocation for new participants. In settings where endpoints are delayed (e.g., survival), the adaptation may not be feasible, or cause significant delays, which subsequently diluting the benefits.

- Inferential challenges:

  Unequal allocation and adaptive design introduce complexities in analysis, requiring specialized methods for estimating treatment effects and controlling type I error rates.[8]

- Potential for overfitting:

  If adaptation occurs too frequently or aggressively, it can lead to unstable estimates or over-reliance on early, potentially spurious signals.

5. Real-World Applications

Bayesian RAR has been used in several notable clinical trials:

- The I-SPY 2 trial in breast cancer employs an adaptive platform design with Bayesian RAR to evaluate multiple treatments simultaneously, adjusting allocation based on biomarker signatures.[9]
- The BATTLE-1 trial was designed to assign more patients to the more promising treatment arms based on data accumulated in the trial up until that time according to each patient's biomarker profile.[10]

These examples demonstrate the versatility of Bayesian RAR across diverse therapeutic areas and trial settings.

6. Future Directions

Bayesian RAR is poised to become increasingly important in the era of personalized medicine, platform trials, and precision therapeutics. As computational tools improve and electronic data capture becomes more seamless, the logistical barriers to adaptive implementation are diminishing.

Integration with machine learning, real-world data, and artificial intelligence may further enhance the responsiveness and utility of Bayesian adaptive trials. Nonetheless, transparency in design, rigorous simulation-based validation, and clear communication with stakeholders remain essential.

7. Conclusion

Bayesian RAR represents a promising evolution in clinical trial design. By integrating ongoing learning with patient-centered ethics, it offers a compelling alternative to traditional fixed designs. While challenges remain, particularly around complexity, delayed outcomes, and regulatory expectations, ongoing methodological advances and real-world experience continue to strengthen the case for wider adoption. In an increasingly dynamic therapeutic landscape, Bayesian RAR offers a framework not only for evaluating treatments more efficiently but also for doing so in a way that better respects the interests of patients and clinicians alike.

Article citation: Yang S, Berdine G. Bayesian response adaptive randomization. The Southwest Journal of Medicine 2025;13(56):56–61
From: Department of Biostatistics, Pennington Biomedical Research Center, Baton Rouge, LA; Department of Internal Medicine (GB), Texas Tech University Health Sciences Center, Lubbock, TX
Submitted: 7/14/2025
Accepted: 7/16/2025
Conflicts of interest: none

References

1. Yang S, Berdine G. Randomization in clinical trials. *The Southwest Respiratory and Critical Care Chronicles* 2022; 10(43):48–51. https://doi.org/10.12746/swrccc.v10i43.1039
2. Robertson DS, Lee KM, López-Kolkovska BC, et al. Response-adaptive randomization in clinical trials: from myths to practical considerations. Stat Sci 2023 May;38(2): 185–208. doi: 10.1214/22-STS865.
3. May Lee K, Lee JJ. Evaluating Bayesian adaptive randomization procedures with adaptive clip methods for multi-arm trials. Stat Methods Med Res 2021 May;30(5):1273–87. doi: 10.1177/0962280221995961.

4. Viele K, Saville BR, McGlothlin A, et al. Comparison of response adaptive randomization features in multiarm clinical trials with control. Pharm Stat 2020 Sep;19(5):602–12. doi: 10.1002/pst.2015.
5. Chongwe G, Ali J, Kaye DK, et al. Ethics of adaptive designs for randomized controlled trials. Ethics Hum Res 2023 Sep–Oct;45(5):2–14. doi: 10.1002/eahr.500178. Erratum in: Ethics Hum Res 2024 Jan–Feb;46(1):13. doi: 10.1002/eahr.500200.
6. Wathen JK, Thall PF. A simulation study of outcome adaptive randomization in multi-arm clinical trials. Clin Trials 2017 Oct;14(5):432–40. doi: 10.1177/1740774517692302.
7. Jiang F, Jack Lee J, Müller P. A Bayesian decision-theoretic sequential response-adaptive randomization design. Stat Med 2013 May 30;32(12):1975–94. doi: 10.1002/sim.5735.
8. Thall P, Fox P, Wathen J. Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. Ann Oncol 2015 Aug;26(8):1621–8. doi: 10.1093/annonc/mdv238
9. Wang H, Yee D. I-SPY 2: a neoadjuvant adaptive clinical trial designed to improve outcomes in high-risk breast cancer. Curr Breast Cancer Rep 2019 Dec;11(4):303–10. doi: 10.1007/s12609-019-00334-2.
10. Liu S, Lee JJ. An overview of the design and conduct of the BATTLE trials. Chin Clin Oncol 2015;4(3):33. doi: 10.3978/j.issn.2304-3865.2015.06.07